Forecasting the quality of water-suppressed ^{1}H MR spectra based on a single-shot water scan

S P Kyathanahally[1, 2] and R Kreis[1]

1. Department of Clinical Research and Institute of Diagnostic, Interventional and Pediatric Radiology, University of Bern, Bern, Switzerland
2. Graduate School for Cellular and Biomedical Sciences; University of Bern, Bern, Switzerland

Running Head: *Forecasting the quality of ^{1}H MR spectra*

Manuscript Word Count: 4727 words

Corresponding author:

Prof. Dr. sc. nat. Roland Kreis,
AMSM, Pavilion 52A, Inselspital,
CH-3010 Bern, Switzerland
Tel: +41-31-632 8174, Fax: +41-31-632 0580
Email: roland.kreis@insel.ch

ABSTRACT

Purpose: To investigate whether an initial non-water-suppressed acquisition that provides information about the signal-to-noise ratio and linewidth is enough to forecast the maximally achievable final spectral quality and thus inform the operator whether the foreseen number of averages and achieved field homogeneity is adequate.

Methods: A large range of spectra with varying signal-to-noise and linewidth was simulated and fitted with popular fitting programs to determine the dependence of fitting errors on linewidth and signal-to-noise. A tool to forecast variance based on a single acquisition was developed and its performance evaluated on simulated and in vivo data obtained at 3 T from various brain regions and acquisition settings.

Results: A strong correlation to real uncertainties in estimated metabolite contents was found for the forecast values and the Cramer-Rao lower bounds obtained from the water-suppressed spectra.

Conclusion: It appears to be possible to forecast the best-case errors associated with specific metabolites to be found in model fits of water-suppressed spectra based on a single water scan. Thus, non-specialist operators will be able to judge ahead of time whether the planned acquisition can possibly be of sufficient quality to answer the targeted clinical question or whether it needs more averages or improved shimming.

Keywords: MR spectroscopy; brain; quantification error; SNR, LW, quality

List of Abbreviations

Asp, Aspartate

Cho, Choline

Cho_{tot}, total Choline (GPC + PCho)

Cr, Creatine

Cr_{tot}, total creatine (Cr + PCr)

GABA, γ-aminobutyric acid

Glc, Glucose

Gln, Glutamine

Glu, Glutamate

Glx, Glutamate + Glutamine

GPC, Glycerophosphocholine

GSH, Glutathione

Gly, Glycine

HLSVD, Hankel-Lanczos Singular Value Decomposition

jMRUI, Java version of the magnetic resonance user interface

Lac, Lactate

mI, Myo-inositol

mI_{tot}, Myo-inositol + Glycine

NAA, N-acetylaspartate

NAAG, N-acetylaspartyl-glutamate

NAA_{tot}, total NAA (NAA + NAAG)

nWS, Non-water-suppressed

PCho, Phosphocholine

PCr, Phosphocreatine

PRESS, Point resolved spectroscopy

QUEST, Quantitation based on quantum estimation

sI, Scyllo-inositol

SD, Standard deviation

SNR, Signal-to-noise ratio

SVS, Single voxel spectroscopy

Tau, Taurine

VESPA, Versatile simulation- pulses and analysis

VOI, Volume of interest

WS, Water-suppressed

INTRODUCTION

Magnetic resonance spectroscopy (MRS) is a non-invasive method that can provide clinically relevant metabolic information (1). Despite the evidence for the value of MRS in clinical practice and despite of technical improvements, the application of MR spectroscopy is still limited because of its challenging nature of interpretation, expert knowledge requirement (2), but also because of time constraints in the clinical setting.

Relative to MRI, MRS has much lower sensitivity and thus requires much longer acquisition times. In addition, MRS requires a more homogeneous B_0 field than MRI to achieve sufficient separation of metabolite resonances, hence necessitating dedicated shimming for each region of interest (ROI) – again prolonging overall acquisition time.

The quantification of in vivo localized proton MR spectra is affected by the signal-to-noise ratio (SNR), broadened line shapes due to inadequate shimming, presence of a non-flat baseline often related to insufficient water suppression, and inherent limits of the contributing spectral components, including overlapping peaks, and the incomplete prior-knowledge in particular with regard to the macromolecular baseline. From the experimental factors, linewidth (LW) and SNR may serve as the main physical measures to rate the quality of spectra (2). The effects of SNR and LW on metabolite concentrations have been investigated at 1.5 T and 4 T (3,4) to show that as SNR decreases or the linewidth broadens, the quantification accuracy clearly decreases. These results suggest that – within limits – knowledge of SNR and LW alone may serve to forecast the achievable quantification accuracy, if other parameters such as patient motion, scanner drift, degrading effects of water suppression, etc., do not interfere for the longer acquisition of the water-suppressed spectrum. Knowledge of the achievable quantification accuracy before actual acquisition of a water-suppressed spectrum may substantially influence the course of clinical MR examinations.

In present clinical routine, the usual protocol consists of choosing the ROI to acquire the spectrum from, optimizing acquisition settings, including shimming, followed by acquisition of the water-suppressed spectrum and - usually afterwards - acquisition of a single-shot water-unsuppressed spectrum for referencing. If expert knowledge is available at that time, a real time evaluation of spectral quality may influence the decision to acquire more repetitions to improve SNR or to change location to improve the shim. If at this point or later in off-line post-processing data quality is found to be insufficient, the scanner/patient time used (including shimming, voxel

placement, and actual scanning) is wasted and often there is no time and opportunity to re-record the spectrum with better-suited acquisition parameters.

Particularly for focal lesions where the size and location of the ROI is different from case to case, even for the expert, it is not straightforward to forecast how long signals should be averaged and how well the ROI has to be shimmed in order to arrive at an interpretable spectrum for the particular clinical question. The operator would have to take into account factors like the size of the ROI, the quality of the shim (influenced by proximity to skull/bone/air transitions, or presence of blood), the field strength and other hardware specifics. However, achievable SNR and LW can essentially be known from a single water-unsuppressed acquisition and if this reference data is acquired before the lengthy recording of water-suppressed data, the decision to go ahead could be taken on more objective grounds. In addition, using the square root dependence of SNR with regard to the number of averages (2), the water signal could be used as basis to forecast how many signal averages will be needed, to achieve a minimally required quantification precision.

In this study, the effects of SNR and LW on metabolite concentrations at 3 T were modeled and it was investigated whether this information could be used to forecast the best achievable modeling inaccuracies for the metabolites of interest before the relevant water-suppressed data is recorded and thus to influence the acquisition parameters for the subsequent acquisition of the water-suppressed spectrum. The performance of the forecast was investigated using simulated spectra and then verified for in vivo spectra of human brain for all major metabolites, including myo-inositol and glutamate.

METHODS

Simulation

Brain metabolite spectra were simulated in Vespa (5) using spin Hamiltonian parameters of metabolites (6,7) for ideal short and long echo time PRESS localization (echo times (TEs) 30 ms and 136 ms) and scaled to yield a normal brain spectrum based on concentrations and T_2 values from literature (8,9). Simulated metabolites for short TE included aspartate (Asp), γ-aminobutyric acid (GABA), creatine (Cr), glucose (Glc), glutamate (Glu), glutamine (Gln), glycine (Gly), glycerophosphocholine (GPC), glutathione (GSH), lactate (Lac), myo-inositol (mI), N-acetylaspartylglutamate (NAAG), N-acetylaspartate (NAA), phosphocholine (PCho),

phosphocreatine (PCr), scyllo-inositol (sI), taurine (Tau). Asp, GABA, Glc were not included for long TE spectra due to their T_2-attenuated minor contributions. Similarly, a water spectrum was simulated and scaled using a T_2 for water of 80 ms (10) and an assumed water content specified by the user (see below). The linewidths for the simulated base spectra of metabolites and water were 2 and 4 Hz, respectively.

A macromolecular baseline (MMBL) signal was also added for the simulated short TE spectra. The MMBL had been acquired from 10 human subjects in a 2DJ experiment with the metabolite-nulling technique (11) using an inversion time of 900 ms and TEs ranging from 20 to 95 ms. For the short TE baseline spectrum, the macromolecule spectra from TE=25, 30, and 35 ms were averaged to create an approximate TE=30 ms spectrum. The MMBL spectrum was preprocessed to remove water and metabolite residuals using Hankel-Lanczos Singular Value Decomposition (HLSVD). Then the MMBL spectrum was modeled as sum of Lorentzians by HLSVD with restraints on damping constants to obtain a noiseless representation of the MMBL spectrum. The linewidths of the base components of this MMBL spectrum were then adjusted to match a metabolite spectrum of 2 Hz width.

These in silico spectra (water and metabolites) were used to generate spectra with different LW. First, the water spectrum was apodized to feature LWs ranging from 5 to 15 Hz in 1 Hz increment, using different Gaussian damping factors. The same damping factors were then applied to the corresponding metabolite spectra.

For each LW, 100 spectra of different noise realizations were created to obtain data sets at 18 SNRs ranging from 5 to 300 (5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 250, 300), where SNR was defined in time domain (TD) as amplitude of the first point of the overall metabolite signal divided by the standard deviation of the noise in the TD signal. In total, 19800 spectra were created for both short and long TE.

Quantification

Simulated spectra were quantified using the two most widely used fitting packages (LCModel (12) and jMRUI (QUEST) (13)). QUEST is a pure TD method whereas LCModel is using the frequency domain (FD) for least squares minimization (14).

LCModel works fully automatic and operator-independent to decompose an in vivo spectrum into a linear combination of model spectra provided as prior knowledge. It also estimates a flexible lineshape model and a spline baseline that accounts for model imperfections (12,15). In this study, LCModel was used in the standard configuration with the analyzing window from 0.2 to 4 ppm.

QUEST also quantifies spectra by utilizing model spectra as prior knowledge. To accommodate a baseline it uses a three step procedure. It is assumed that the background signals have short T_2 and therefore truncating the initial points would separate the background from the metabolite signals (16,17). So first, initial points are truncated and the metabolite parameters are estimated on the truncated fid. The fitted truncated signal is then back-projected to the time origin and subtracted from the original signal to get the noisy background signal. In the second step, the noisy background signal is modeled by Lorentzian components using SVD. In the third step, the parameterized background signal is either subtracted from the raw signal or – as applied here - used as a further model component (option “InBase”) and a parametric nonlinear least-squares fit is performed in TD. In our study, the linewidth parameter was allowed to range between -10 and +150 Hz while frequency shifts were not allowed, assuming previous overall alignment and no relative shifts of metabolite components. The ‘InBase’ option was chosen with 5 truncation points, which was found optimal to model very broad baseline components for the case where the MMBL are handled by including a MMBL spectrum in the basis set. .

Both methods yield estimates for metabolite content along with Cramer-Rao lower bounds (CRLB) for the fitting errors. CRLB (18) reflect minimal uncertainties for each fit as given by the interdependence of fitting parameters in the model and the experimental noise level associated with the measurement. The CRLB represent the lowest possible error, with overall uncertainties bound to be larger because of additional experimental uncertainties.

Prior knowledge for the quantification included the information describing all resonances present in the simulated spectra. For short TE spectra, the MMBL spectrum was also included as basis spectrum since it has been shown (19) that inclusion of a metabolite-nulled spectrum in the basis set leads to more reliable and accurate metabolite quantification than using the estimated MMBL from LCModel. For QUEST, at both long and short TE, the prior knowledge did not include separate Cr and PCr but rather a sum of Cr and PCr (50:50), in order to obtain an estimate and CRLB of total creatine (Cr_{tot} = Cr +PCr).

The median CRLB, median bias (fit minus true values) and the standard deviations (SD) of metabolite contents across 100 spectra were calculated for each combination of LW and SNR and tabulated for the forecasting tool. The medians were preferred over averages for their insensitivity to single cases where the fit algorithm may not have found the global χ^2 minimum.

The absolute concentrations and CRLB error estimates were calculated with respect to water. LCModel provides the CRLB in percent, so to express them in absolute values; they were multiplied with the estimated concentrations.

Forecast tool

For forecasting metabolite errors in the form of CRLB of fitted spectra for metabolite content based on the water spectrum, first the linewidth of the experimental non-water suppressed spectrum was determined by fitting the signal to a Voigt lineshape. Then the expected SNR of the metabolite spectrum was calculated based on the proposed number of averages, the TE and the water content (WC) of the ROI as indirectly indicated by the user as relative white and gray matter content of the ROI. 75% water content in pure gray matter and 65% in pure white matter were assumed. (T_1-relaxation effects and CSF contributions to the water signal are currently neglected in this prototype tool).

Equation 1 shows how the expected TD SNR of the metabolite spectrum was calculated based on a single acquisition of the water signal.

$$\text{Expected SNR of metabolite TD signal} = \left(\frac{SNR_{water}}{K}\right)\sqrt{averages}/\,WC \qquad [1]$$

where

SNR_{water} is the SNR in the single shot water signal,

$$\text{K} = \frac{initial\ magnitude\ of\ TD\ signal\ of\ simulated\ water}{initial\ magnitude\ of\ TD\ signal\ of\ sum\ of\ simulated\ metabolite\ signals}.$$

and averages stands for the prescribed number of acquisitions for the water-suppressed spectrum.

Using LW and the SNR information, the errors associated with each metabolite concentration estimate were determined by interpolation of the tabulated values.

The whole procedure of forecasting is shown in a flowchart in Figure 1

Performance testing

In-silico

To test how well the forecast works at best, the water spectra simulated with varying LWs and SNRs were treated as experimental spectra to forecast the metabolite errors for the associated metabolite spectra. The differences with respect to CRLB as well as the true variance found in the model fits were investigated as a function of LW and SNR for both short and long TE.

In-vivo

The forecast tool was then applied to in-vivo cases. Twelve healthy subjects (4 women and 8 men) gave informed consent according to the procedure approved by the local ethics committee. The scans were performed on 3-T scanners (Trio and Verio, Syngo MR VB17 Siemens, Erlangen, Germany) using 16 channel phased array receive-only head coils. Each measurement session included basic MRI to define the ROIs for MRS. First- and second-order shims were adjusted using FASTESTMAP (20). Water suppressed, as well as unsuppressed spectra were acquired using a standard PRESS sequence at TE = 30 ms and 136 ms, TR = 3000 ms, number of acquisitions ranging from 16 to 128, ROI size ranging from 1.6 to 9.1 cm^3 (with typical dimensions from 16x10x10 mm^3 to 40x19x12 mm^3) and from different parts of the brain (frontal, occipital, supraventricular) containing both white matter and gray matter.

The in vivo spectra were analyzed using jMRUI and LCModel with parameters set as described for the simulated datasets. For jMRUI, preprocessing included eddy current-correction using the phase information of the water spectrum and removal of residual water using HLSVD decomposition. In LCModel, the option to do eddy-current-correction using the phase information of the water reference was chosen. Residual water was not removed.

RESULTS

Forecasting with LCModel

LCModel was able to quantify all 19800 simulated metabolite spectra at short and long TE. Figure 2 shows representative spectra and fit results from the large set of simulated spectra,

illustrating that LCModel (left panel) did not yield substantial systematic residuals in any of the situations, i.e. neither for a spectrum with substantial noise (SNR 20) in row a, nor the noiseless cases without broadening (row b) or with 10 Hz Gaussian broadening (row c), where LCModel profits from a flexible shape estimation with sufficient freedom to adapt to Gaussian broadening.

Systematic residuals would lead to bias for the estimated metabolite contents. The median bias expressed as percentage of the true concentrations was found to be very small or even negligible for LCModel at short TE with +3%, +2%, +3%, +4%, -3% and <1% for Cr_{tot}, NAA_{tot}, Glu, Glx, mI_{tot} and Cho_{tot}, respectively. Similar results were found at long TE.

Figure 3 and Figure 4 illustrate the database used for developing the forecast tool with estimated CRLB and true variance found when fitting the simulated spectra. The true variance i.e. the SDs of metabolite estimates across the 100 spectra for each particular combination of LW and SNR was compared to the CRLB given by the fitting tool. These figures show the average CRLB (on the left) and true variance in estimated metabolite content (on the right) as a function of LW and SNR for a selection of metabolites in the cases of short (LCModel results in Fig 3a) and long (Fig 4a) TE. As expected, uncertainties of estimated metabolite contents, reflected in CRLB and true variance, generally increase when the lines broaden or SNR decreases. CRLB and true variance show similar values and dependences for LCModel. In general, it was found that the CRLB somewhat overestimate the variance of fitting results. In particular at TE 30 ms, the ratio between CRLB and true SD was between 1.05 and 1.29 for the major metabolites. Looking into the detailed dependence of the comparison between CRLB and true variance on linewidth, we find that e.g. for Cr_{tot}, the ratio increased from 1.1 at FWHM of 5 Hz to 1.3 at 15 Hz.

These sets of CRLB and true variance were then translated into large lookup tables for expected errors based on estimated SNR and LW of the related single shot water signals. Figs 3 and 4 represent these look-up tables. To judge the in silico performance, the mean absolute difference between the forecast error based on the water signal and the actual error found from the fits was evaluated. At short TE, the maximum difference was just 1% for LCModel, such that from the synthetic data it appears that an error forecast based on the water signal seems feasible for many metabolites.

The proposed scheme was verified on in vivo spectra that had SNR and FWHM ranging from 5 to 190 and 5 to 14 Hz, respectively. Example in vivo data is shown in Figure 5 with different LWs, SNRs, voxel location and TEs. The top row shows the initially acquired water scan with

estimated linewidth and SNR. The middle row presents the predicted CRLB values from the look up table and simulated water-suppressed spectra that closely match the corresponding LW and SNR. The bottom row illustrates the subsequently actually acquired in vivo metabolite spectra and the CRLB values obtained for their fits. Figures 6 to 7 document the outcome for all in vivo cases by presenting the correlation between the predicted CRLB based on the water acquisition and those found from the fitting results of the actually recorded WS spectra for some metabolites. In addition, a Bland-Altman plot is added for each correlation to investigate potential systematic deviations. Results are illustrated for short (Figure 6) and long (Figure 7) TE. The coefficients range between 0.73 and 0.91 for the metabolites shown. The Bland-Altman plots show a very small but significant offset for some of the metabolites and no significant bias for others. The confidence intervals for the deviation between forecast and actual errors as given in the plots range between ±0.03 mM (Ch_{tot}) and ±0.3 mM (Glu and Glx) for the case of short TE in LCModel. The situation is similar at long TE.

Forecasting with jmrui/QUEST

The major outcome was similar when using QUEST instead of LCModel, but the fact that QUEST is restricted to Lorentz-broadening lead to larger systematic deviations and a weaker performance for the CRLB forecasts. The major differences to LCModel are presented in detail in the following.

QUEST was able to quantify all long TE spectra, but failed for 2.6% of the short TE cases. Fit results for jMRUI's QUEST (Figure 2 right panel) are similar to those of LCModel in cases with substantial noise (row a), and also for the noiseless case without broadening (row 2b). However, residuals are much larger for QUEST for a case with 10 Hz Gaussian broadening (row 2c), where the Voigt lines are dominated by the Gaussian broadening but are fitted with a purely Lorentzian lineshape model.

These systematic residuals are reflected in substantial bias for the estimated metabolite contents in the synthetic spectra for some of the investigated cases. The median bias at short TE was found to be much larger for QUEST than for LCModel with values of +32% and +29% for Cr_{tot} and Glu compared to +3% for both cases when using LCModel.

For QUEST, the comparison between average CRLB (on the left in Figs 3 and 4) and true variance (on the right in Figs 3 and 4) as a function of LW and SNR is hampered by the fact that

QUEST does not provide CRLB estimates for sums of metabolites, but - where available - the differences between true variance and estimated CRLB are larger for QUEST than for LCModel. As with LCModel, it was found that the CRLB overestimate the variance of fitting results systematically. However for QUEST, even though the noise was calculated using the original data and not the residues, the ratio between CRLB and true SD was larger and in the range of 2.1 to 3 compared to below 1.3 for LCModel.

Looking at the in silico performance, the mean absolute difference between the forecast error and the actual error was found to be somewhat larger for QUEST than LCModel (maximum difference of 8 vs. 1%), but still small enough to promise valid results also for QUEST.

Results for the in-vivo performance of error forecasts using QUEST are illustrated for short TE (Figure 8) and long TE (Figure 9). For QUEST, correlation coefficients range between 0.78 and 0.89, similar to the values for LCModel while Bland-Altman plots similarly show small but significant offsets for some metabolites. The confidence intervals for these deviations (see Figs 8-9) are again somewhat larger for QUEST.

DISCUSSION

A water-unsuppressed signal is nearly always recorded in clinical MRS investigations and can be used as a reference for lineshape correction and as an internal reference for absolute metabolite quantitation. Here, we show that it is also possible to use such a single-shot water signal to forecast an estimate for the quantification error that will be associated with the fitting result for most relevant metabolites from the subsequently recorded water-suppressed spectrum. The forecast is based on the linewidth and SNR of the water data, as well as the planned number of acquisitions.

Clinical MRS exams are most often performed by non-expert personnel. They cannot judge the quality of the acquired water scan or spectrum and thus whether or not any targeted question about potential changes in metabolite content could possibly be answered given the foreseen or even the already acquired spectrum. The possibility to forecast fitting errors would allow defining minimal quality prerequisites to be fulfilled for an MR spectrum to be recorded under the planned acquisition settings. In particular, this may allow the non-specialist (and even the experienced user) to better judge in real time, i.e. with the patient still in the scanner, how many acquisitions should be prescribed for a certain clinical question, or whether under the given

circumstances the ROI should be increased or possibly moved away from areas of potential susceptibility problems to obtain narrower linewidths. If quality screening is only performed after acquisition of the WS spectrum, time will often not allow for improved or prolonged spectral acquisitions and of course if spectral processing is done off-line, enhanced or longer acquisitions are not feasible.

Not very surprisingly, the forecast worked very well for the synthetic cases, also when comparing the observed variance from fitting 100 different spectra per case with the forecast CRLB. But as evident from Figs. 6 to 9, the error forecast showed very good accuracy also in vivo. For reasons of limited space, the results were only presented for the clinically most interesting metabolites (and for this reason primarily for the more accurately defined sums of metabolites with a similar spectrum) at both echo times, but results are similar for the other metabolites and there is no reason why it could not be used for any metabolite or other acquisition condition, if needed.

However, the value of the forecast is limited in several aspects. 1) The forecast errors provide a best case scenario. If the patient moves, or acquisition conditions are not constant for other reasons, the forecast may be incorrect. Also, if the WS spectrum should be plagued by artifacts (2), the best case scenario would not apply – thus the proposed tool should only be seen as a minimal quality requirement, not as a guarantee for eventual success of the MRS exam. In fact, it should be seen as a warning system for likely failure, not as an indicator of probable success. 2) The proposed forecast tool should be implemented in a real time version, best on the scanner console or at least with automatic transfer and evaluation on a satellite station next to the scanner, since otherwise the time benefit would be wasted. 3) For each clinical question a threshold uncertainty has to be predefined requiring the experience of previous investigations, but this can be defined based on literature data. 4) The currently presented forecast is based on spectra that are composed of those metabolites that occur in healthy brain. Interference with other spectral elements, e.g. lipid resonances, may be investigated in addition.

Further limitations concern general issues of fitting MR spectra, where it is known that results will depend on the fitting and processing routines used and in particular the prior knowledge constraints included. Hence the forecast errors will only be applicable if an evaluation similar to the one used to define the forecast tool is being used. This has become evident from the comparison of the LCModel and QUEST results, as described above. In particular, baseline and lineshape issues may well modify the results and the effective fitting error drastically. The baseline signals pose a challenge for obtaining reliable estimates, because they may have

unknown shapes and intensities and may include lipids. The macromolecules and lipids also vary based on pathologies (21–25). Accurate estimation of macromolecules is needed for reliable quantification of metabolites. For the synthetic spectra used in this study, this was confirmed when comparing different evaluation schemes for both LCModel and QUEST, when we did or didn't use a macromolecular baseline as part of the basis set (results not shown). For QUEST, the results depended heavily on the algorithm of baseline incorporation (e.g. number of initial data points to use for baseline definition). In particular, the baseline is a serious problem when the spectral resolution is low. Over- or underestimation of the macromolecule signal may strongly influence content and CRLB of the overlapping metabolites with broad spectral patterns. It should also be noted that the forecast errors as defined in the present set up would not be correct in cases of pathology, when other base spectra would be introduced, or lipids would have to be included in the fit.

The lineshape due to the remaining field inhomogeneity after shimming is often non-Lorentzian and thus fitting with pure Lorentzians easily leads to systematic errors (26). Experimental lineshapes are more precisely modeled by a Voigt function (multiplication of Lorentzian and Gaussian damping). Hence, we used the Voigt lineshape to fit the water spectrum for the determination of the water linewidth. If the true LWs deviate from this model, this might impact on the forecast. This is currently not considered, but a check of lineshape could be included in the forecast. The fact that QUEST in its current form did not allow for use of Voigt lines or general lineshapes certainly is the reason why it showed considerable bias. In contrast, LCModel uses a regularized lineshape model that accounts for the unknown lineshapes thus reducing effects of lineshape on bias. The effect of the used lineshape is evident in Fig. 2.

A minor issue that has not been treated in depth so far is the exact water content of the investigated ROI. At present, this is included indirectly with the user describing the ROI content in terms of white/grey matter percentage. Using a proton density map – possibly augmented with automatic tissue segmentation, this could be circumvented, such that also for cases of cysts and necrosis proper scaling could be achieved (27).

The fitting results from the vast set of simulated spectra with differing SNR and LW also yielded insight into the performance of LCModel and QUEST as well as into general dependence of fit results on LW and SNR. Figs 3 -4 illustrate this dependence with the diagrams on true variance (right side panels). The outcome is similar to earlier investigations (3,4). It should be noted that these diagrams depend strongly on whether the SNR is defined in time or frequency domain. Here we have used the TD definition, because in this way SNR does not depend on the

linewidth. With the FD definition of SNR, constant SNR with variable linewidth implicate differing acquisition times. Thus, in the presented figures, one can observe the expected larger variance for an increase in linewidth whereas if the SNR is defined in frequency domain, one may even detect a decrease in the measurement error with increased linewidth just because this mixes the effect of linewidth and length of acquisition time.

From Figs 3 and 4, it seems that the absolute error for long TE is generally lower than for short TE. This is only true for identical SNR, but to achieve similar SNR in these particular cases, it would roughly take at least twice as long at the longer TE. Hence, for equal acquisition times, the fitting errors for those compounds that can be measured at both echo times would probably be about the same.

As an outlook, the performance of this forecast tool might be optimized by use of actual in vivo data as part of the look-up table, where one might also improve the performance by including lineshape characteristics other than the simple LW and possibly machine learning algorithms for the predictor.

In conclusion, our results indicate that it is possible to forecast the errors associated with specific metabolites to be found in model fits of WS spectra based on a single shot water signal. Thus, the clinical MRS user will be able to judge ahead of time whether the planned acquisition can possibly be of sufficient quality to answer the question at hand or whether it needs more averages or improved shimming.

Acknowledgement

This research was carried out in the framework of the European Marie-Curie Initial Training Network, 'TRANSACT', PITN-GA-2012-316679, 2013-2017 and also supported by the Swiss National Science Foundation (#320030_156952).

REFERENCES

1. Oz G, Alger JR, Barker PB, et al. Clinical proton MR spectroscopy in central nervous system disorders. Radiology 2014;270:658–679. doi: 10.1148/radiol.13130531.

2. Kreis R. Issues of spectral quality in clinical 1H-magnetic resonance spectroscopy and a gallery of artifacts. NMR Biomed. 2004;17:361–381. doi: 10.1002/nbm.891.

3. Kanowski M, Kaufmann J, Braun J, Bernarding J, Tempelmann C. Quantitation of simulated short echo time 1H human brain spectra by LCModel and AMARES. Magn Reson Med. 2004;51:904–912. doi: 10.1002/mrm.20063.

4. Bartha R. Effect of signal-to-noise ratio and spectral linewidth on metabolite quantification at 4 T. NMR Biomed. 2007;20:512–521. doi: 10.1002/nbm.1122.

5. Soher B, Semanchuk D, Todd D, Steinberg J, Young K. Vespa: Integrated applications for RF pulse design, spectral simulation and MRS data analysis. In: Proceedings of the 19th Annual Meeting ISMRM. Montreal,Canada; 2011. p. 1410.

6. Govindaraju V, Young K, Maudsley AA. Proton NMR chemical shifts and coupling constants for brain metabolites. NMR Biomed 2000;13:129–153.

7. Govind V, Young K, Maudsley AA. Corrigendum: Proton NMR chemical shifts and coupling constants for brain metabolites. Govindaraju V, Young K, Maudsley AA. NMR Biomed . 2000; 13: 129-153. NMR Biomed. 2015;28:923–924. doi: 10.1002/nbm.3336.

8. Mekle R, Mlynárik V, Gambarota G, Hergt M, Krueger G, Gruetter R. MR spectroscopy of the human brain with enhanced signal intensity at ultrashort echo times on a clinical platform at 3T and 7T. Magn Reson Med. 2009;61:1279–1285. doi: 10.1002/mrm.21961.

9. Mlynárik V, Gruber S, Moser E. Proton T(1) and T(2) relaxation times of human brain metabolites at 3 Tesla. NMR Biomed. 2001;14:325–331. doi: 10.1002/nbm.713.

10. Provencher S. LCModel & LCMgui User's Manual. 2015:1–184.

11. Behar KL, Rothman DL, Spencer DD, Petroff OAC. Analysis of macromolecule resonances in 1H NMR spectra of human brain. Magn Reson Med. 1994;32:294–302. doi: 10.1002/mrm.1910320304.

12. Provencher SW. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. Magn Reson Med. 1993;30:672–679. doi: 10.1002/mrm.1910300604.

13. Naressi A, Couturier C, Castang I, Beer R De, Graveron-Demilly D. Java-based graphical user interface for MRUI , a software package for quantitation of in vivo/medical magnetic resonance spectroscopy signals. Comput Biol Med. 2001;31:269–286. doi: 10.1016/S0010-4825(01)00006-3.

14. Poullet J-B, Sima DM, Van Huffel S. MRS signal quantitation: a review of time- and

frequency-domain methods. J Magn Reson. 2008;195:134–144. doi: 10.1016/j.jmr.2008.09.005.

15. Provencher SW. Automatic quantitation of localized in vivo 1H spectra with LCModel. NMR Biomed. 2001;14:260–264. doi: 10.1002/nbm.698.

16. Ratiney H, Sdika M, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D. Time-domain semi-parametric estimation based on a metabolite basis set. NMR Biomed. 2005;18:1–13. doi: 10.1002/nbm.895.

17. Ratiney H, Coenradie Y, Cavassila S, Van Ormondt D, Graveron-Demilly D. Time-domain quantitation of 1H short echo-time signals: Background accommodation. Magn Reson Mater Phy. 2004;16:284–296. doi: 10.1007/s10334-004-0037-9.

18. Cavassila S, Deval S, Huegen C, Van Ormondt D, Graveron-Demilly D. Cramer-Rao bounds: An evaluation tool for quantitation. NMR Biomed. 2001;14:278–283. doi: 10.1002/nbm.701.

19. Cudalbu C, Mlynárik V, Xin L, Gruetter R. Quantification of in vivo short echo-time proton magnetic resonance spectra at 14.1 T using two different approaches of modelling the macromolecule spectrum. Meas Sci Technol. 2009;20:104034.

20. Gruetter R. Automatic, localized in vivo adjustment of all first-and second-order shim coils. Magn Reson Med. 1993;29:804–811. doi: 10.1002/mrm.1910290613.

21. Mader I, Seeger U, Weissert R, Klose U, Naegele T, Melms A, Grodd W. Proton MR spectroscopy with metabolite-nulling reveals elevated macromolecules in acute multiple sclerosis. Brain 2001;124:953–961. doi: 10.1093/brain/124.5.953.

22. Seeger U, Klose U, Mader I, Grodd W, Nägele T. Parameterized evaluation of macromolecules and lipids in proton MR spectroscopy of brain diseases. Magn Reson Med. 2003;49:19–28. doi: 10.1002/mrm.10332.

23. Howe FA, Barton SJ, Cudlip SA, et al. Metabolic profiles of human brain tumors using quantitative in vivo 1H magnetic resonance spectroscopy. Magn Reson Med. 2003;49:223–232. doi: 10.1002/mrm.10367.

24. Graham GD, Hwang J, Rothman DL, Prichard JW. Spectroscopic assessment of alterations in macromolecule and small-molecule metabolites in human brain after stroke. Stroke. 2001;32:2797–2802.

25. Saunders DE, Howe FA, Van Den Boogaart A, Griffiths JR, Brown MM. Discrimination of metabolite from lipid and macromolecule resonances in cerebral infarction in humans using short echo proton spectroscopy. J Magn Reson Imaging 1997;7:1116–1121. doi: 10.1002/jmri.1880070626.

26. Marshall I, Higinbotham J, Bruce SD, Freise A. Use of Voigt line shape for quantification of in vivo 1H spectra. Magn Reson Med. 1997;37:651–657.

27. Tisell A, Leinhard OD, Warntjes JBM, Lundberg P. Procedure for quantitative 1H magnetic resonance spectroscopy and tissue characterization of human brain tissue based on the use of quantitative magnetic resonance imaging. Magn Reson Med. 2013;70:905–915.

Figure Captions

Figure 1: Flowchart illustrating the forecasting procedure. The steps marked in blue are calculated by the toolbox while the steps marked in pink represent input by the user.

Figure 2: Illustration of the fits in LCModel and QUEST for simulated PRESS ^{1}H-MR spectra at TE=30 ms. a) FWHM = 7 Hz, SNR = 20; b) FWHM = 2 Hz (Lorentzian lineshape), no noise added; c) FWHM = 10 Hz (Voigt lineshape), no noise added. The residuals (original spectrum minus fit) are shown above the spectra. For QUEST, the original spectra and their corresponding fits were zero-filled before plotting for better comparison, since zero-filling is part of the LCModel data treatment.

Figure 3: Mean variance for some estimated metabolite content obtained for simulated spectra at short TE when fit with LCModel (a) and QUEST (b). For both cases, the mean CRLB is shown on the left and standard deviations of the determined metabolite contents as obtained over 100 different noise realizations on the right.

Since QUEST does not provide error estimates of summed metabolite areas, some diagrams are missing. It should also be noted, that for Cr_{tot} (where we do indicate CRLB values even in the case of QUEST), we had included a base spectrum for a 50/50 mixture of Cr and PCr, rather than individual spectra in the case of QUEST.

Abbreviations: Cho_{tot} = Glycerophosphocholine + Phosphocholine; Cr = Creatine; CRLB = Cramer-Rao-lower bound; Cr_{tot} = Creatine + Phosphocreatine; Glx = Glutamate + Glutamine; mI_{tot} = myo-Inositol + Glyine; NAA_{tot} = N-acetylaspartate + N-acetylaspartyl-glutamate; PCr = Phosphocreatine; TrueSD= true standard deviation.

Figure 4: Mean variance for some estimated metabolite content obtained for simulated spectra at long TE when fit with LCModel (a) and QUEST (b). For both cases, the mean CRLB is shown on the left and standard deviations of the determined metabolite contents as obtained over 100 different noise realizations on the right.

Since QUEST does not provide error estimates of summed metabolite areas, some diagrams are missing. It should also be noted, that for Cr_{tot} (where we do indicate CRLB values even in the case of QUEST), we had included a base spectrum for a 50/50 mixture of Cr and PCr, rather than individual spectra in the case of QUEST. For Abbreviations, see Fig. 2.

Figure 5: Illustration of sample in vivo data with different LWs, SNRs, voxel location and TEs. The top row shows the single shot water scan. The water line width is calculated, and based on the TE and voxel composition, the SNR is estimated. The middle row shows simulated water suppressed spectra representing the estimated SNR and LW as well as the predicted CRLB from the look-up table. Note that the simulated spectrum is just for illustration and that the SNR and LW do not match exactly. The bottom row shows the subsequently acquired in vivo metabolite spectra and the corresponding CRLB obtained from fitting these spectra.

Figure 6: Illustration of the in vivo performance of the proposed error forecast method for the case of LCModel at short TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 7: Illustration of the in vivo performance of the proposed error forecast method for the case of LCModel at long TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 8: Illustration of the in vivo performance of the proposed error forecast method for the case of QUEST at short TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 9: Illustration of the in vivo performance of the proposed error forecast method for the case of QUEST at long TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

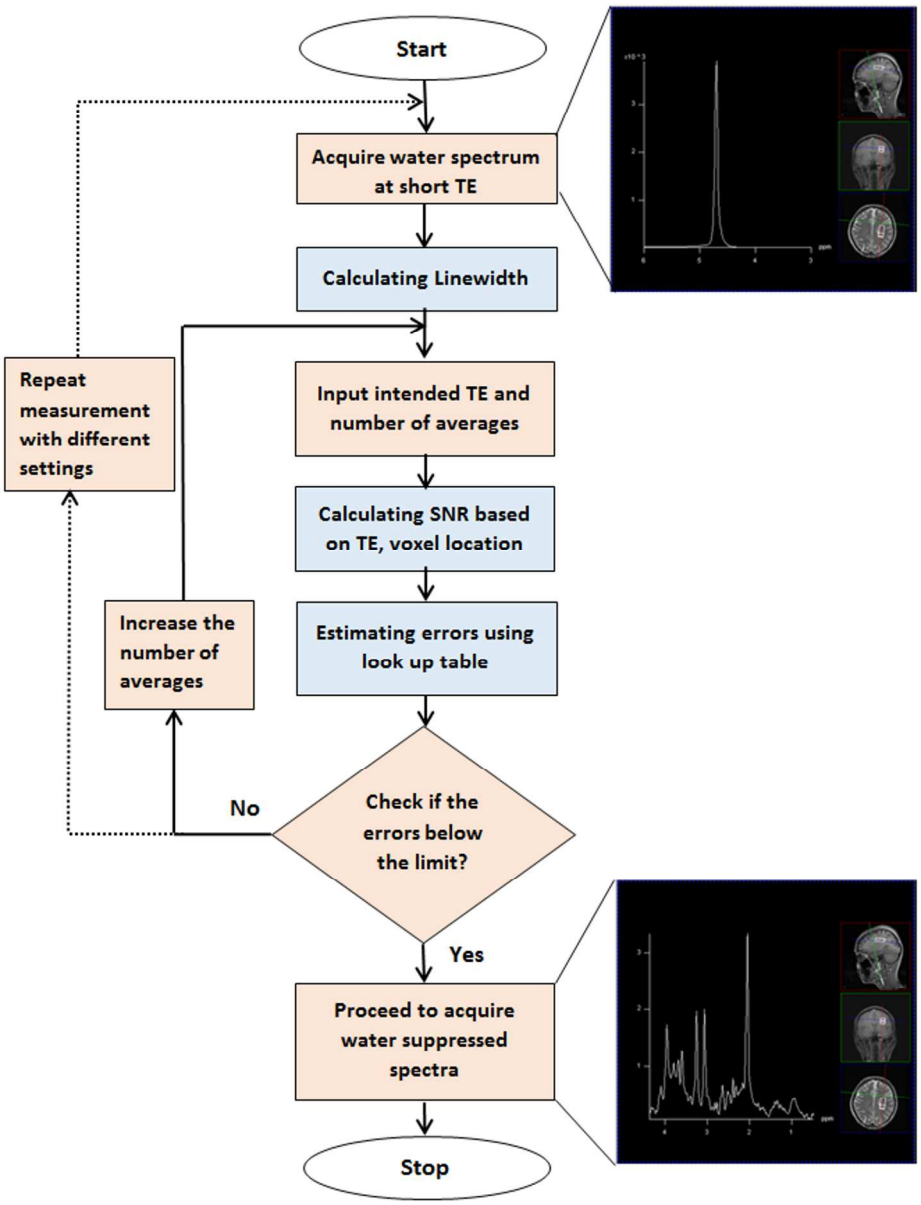

Figure 1: Flowchart illustrating the forecasting procedure. The steps marked in blue are calculated by the toolbox while the steps marked in pink represent input by the user.
Figure 1
229x299mm (300 x 300 DPI)

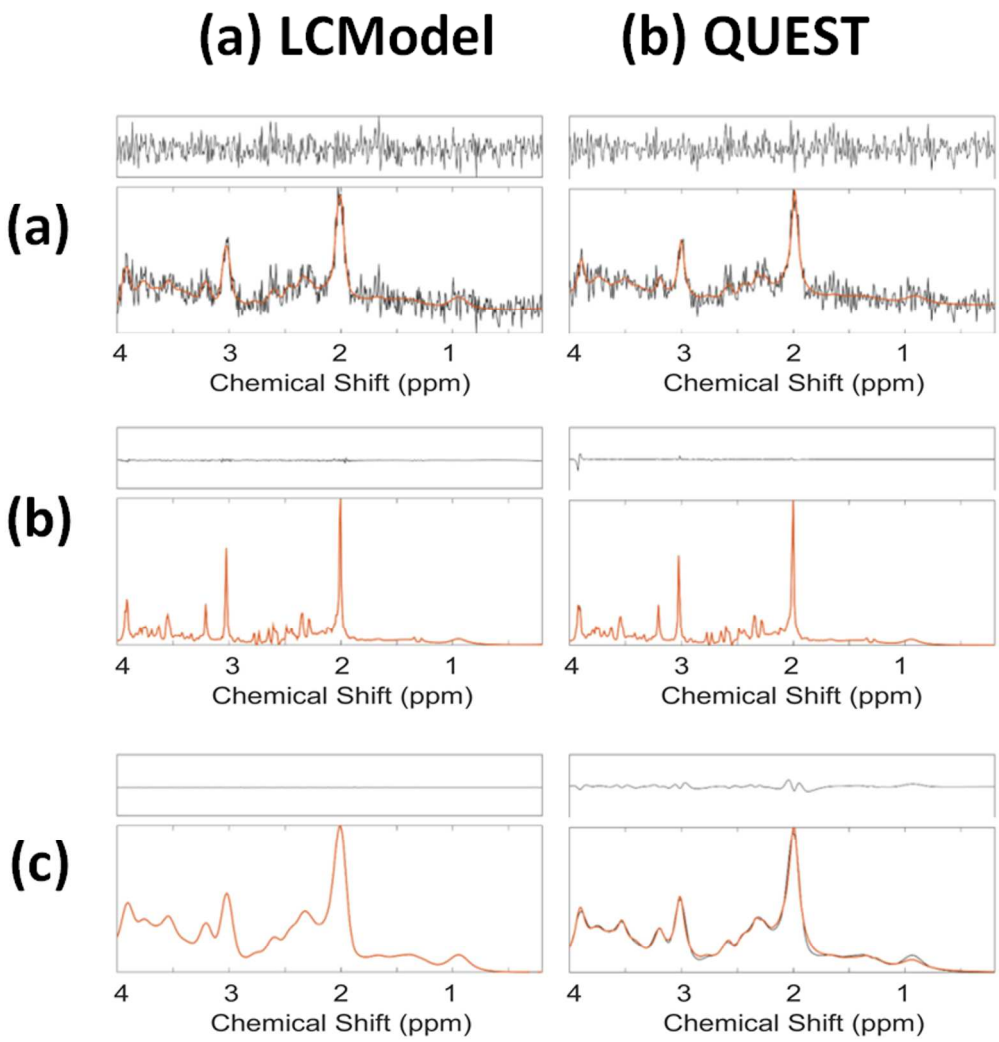

Figure 2: Illustration of the fits in LCModel and QUEST for simulated PRESS 1H-MR spectra at TE=30 ms. a) FWHM = 7 Hz, SNR = 20; b) FWHM = 2 Hz (Lorentzian lineshape), no noise added; c) FWHM = 10 Hz (Voigt lineshape), no noise added. The residuals (original spectrum minus fit) are shown above the spectra. For QUEST, the original spectra and their corresponding fits were zero-filled before plotting for better comparison, since zero-filling is part of the LCModel data treatment.
Figure 2
91x96mm (600 x 600 DPI)

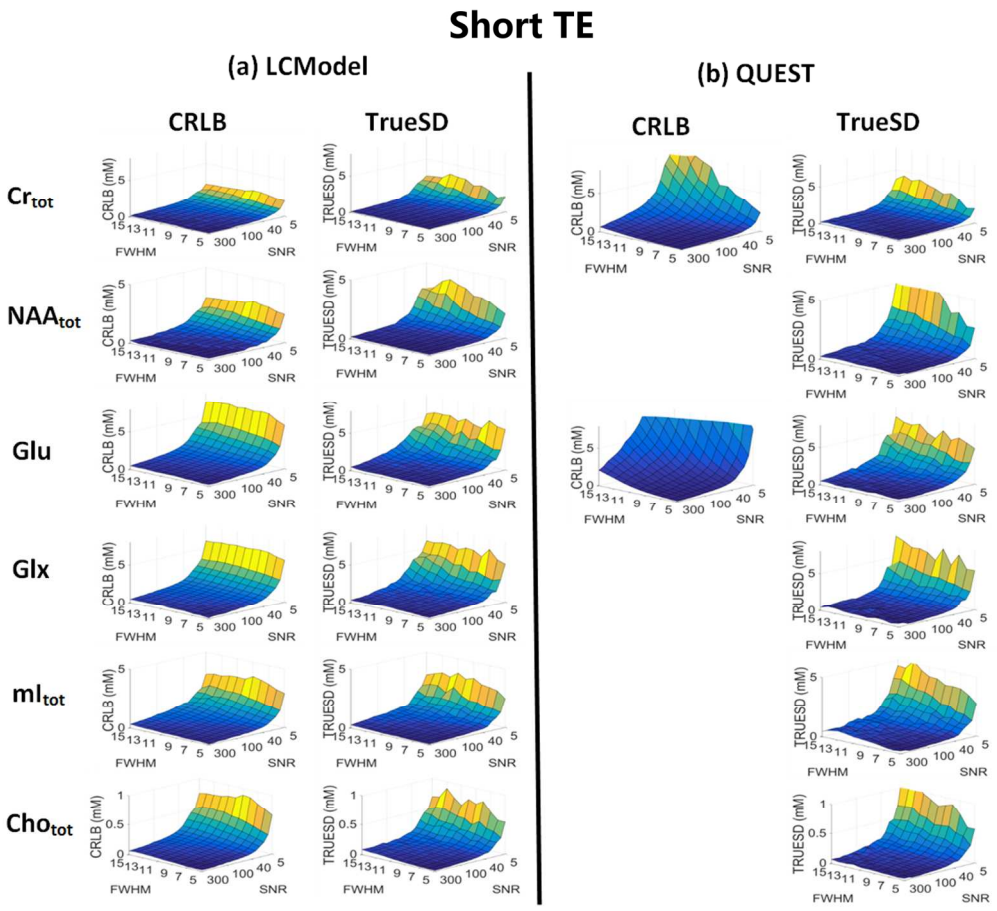

Figure 3: Mean variance for some estimated metabolite content obtained for simulated spectra at short TE when fit with LCModel (a) and QUEST (b). For both cases, the mean CRLB is shown on the left and standard deviations of the determined metabolite contents as obtained over 100 different noise realizations on the right.\r\nSince QUEST does not provide error estimates of summed metabolite areas, some diagrams are missing. It should also be noted, that for Crtot (where we do indicate CRLB values even in the case of QUEST), we had included a base spectrum for a 50/50 mixture of Cr and PCr, rather than individual spectra in the case of QUEST.\r\nAbbreviations: Chotot = Glycerophosphocholine + Phosphocholine; Cr = Creatine; CRLB = Cramer-Rao-lower bound; Crtot = Creatine + Phosphocreatine; Glx = Glutamate + Glutamine; mItot = myo-Inositol + Glyine; NAAtot = N-acetylaspartate + N-acetylaspartyl-glutamate; PCr = Phosphocreatine; TrueSD= true standard deviation.

Figure 3
160x145mm (300 x 300 DPI)

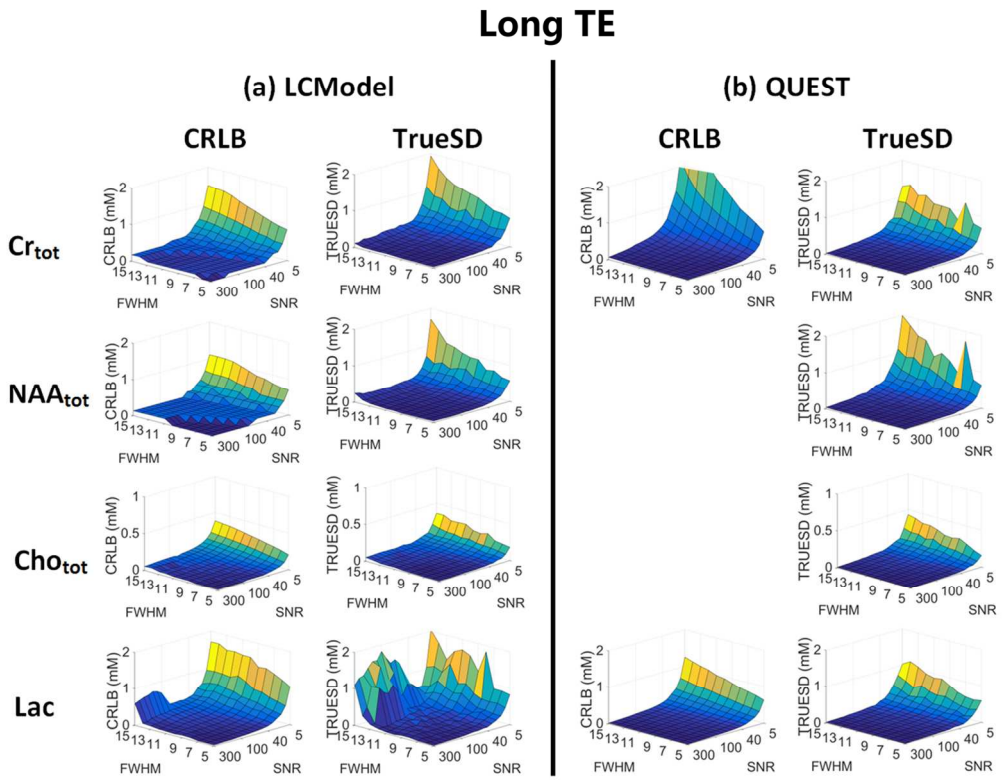

Figure 4: Mean variance for some estimated metabolite content obtained for simulated spectra at long TE when fit with LCModel (a) and QUEST (b). For both cases, the mean CRLB is shown on the left and standard deviations of the determined metabolite contents as obtained over 100 different noise realizations on the right.\r\nSince QUEST does not provide error estimates of summed metabolite areas, some diagrams are missing. It should also be noted, that for Crtot (where we do indicate CRLB values even in the case of QUEST), we had included a base spectrum for a 50/50 mixture of Cr and PCr, rather than individual spectra in the case of QUEST. For Abbreviations, see Fig. 2.

Figure 4
139x111mm (300 x 300 DPI)

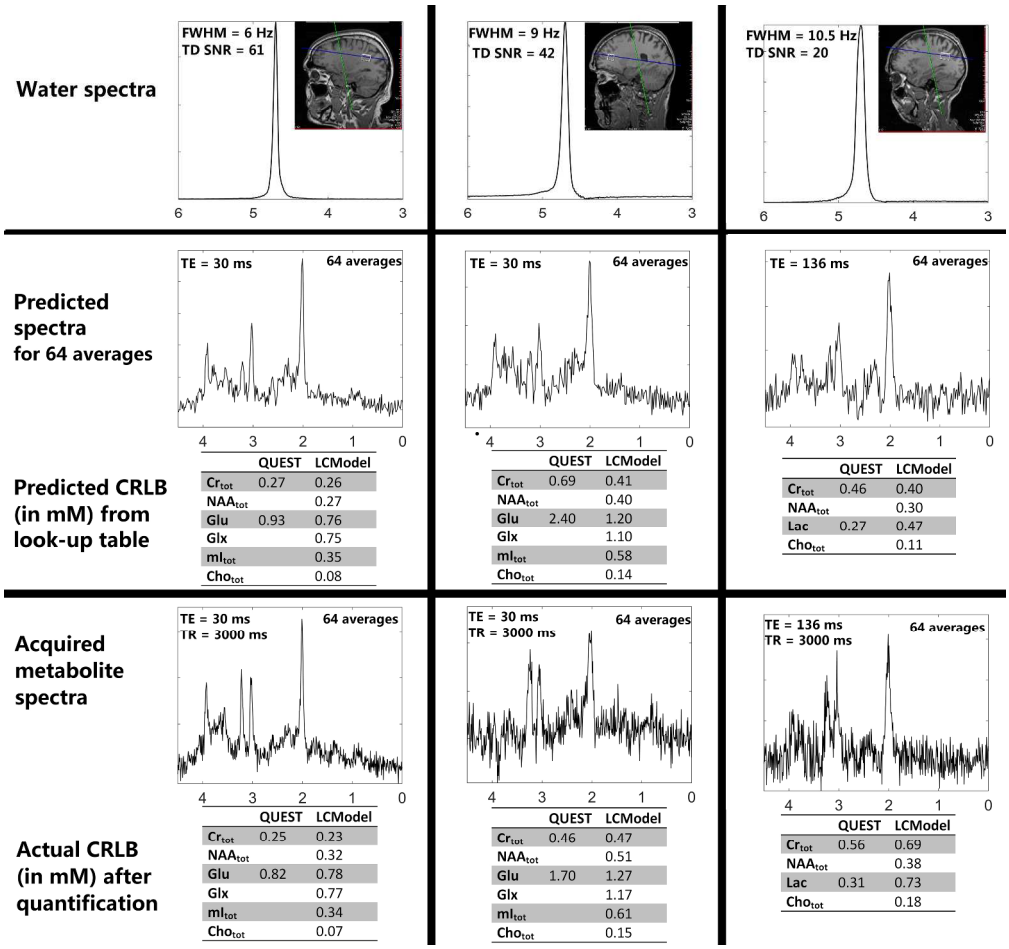

Figure 5: Illustration of sample in vivo data with different LWs, SNRs, voxel location and TEs. The top row shows the single shot water scan. The water line width is calculated, and based on the TE and voxel composition, the SNR is estimated. The middle row shows simulated water suppressed spectra representing the estimated SNR and LW as well as the predicted CRLB from the look-up table. Note that the simulated spectrum is just for illustration and that the SNR and LW do not match exactly. The bottom row shows the subsequently acquired in vivo metabolite spectra and the corresponding CRLB obtained from fitting these spectra.

Figure 5
164x154mm (600 x 600 DPI)

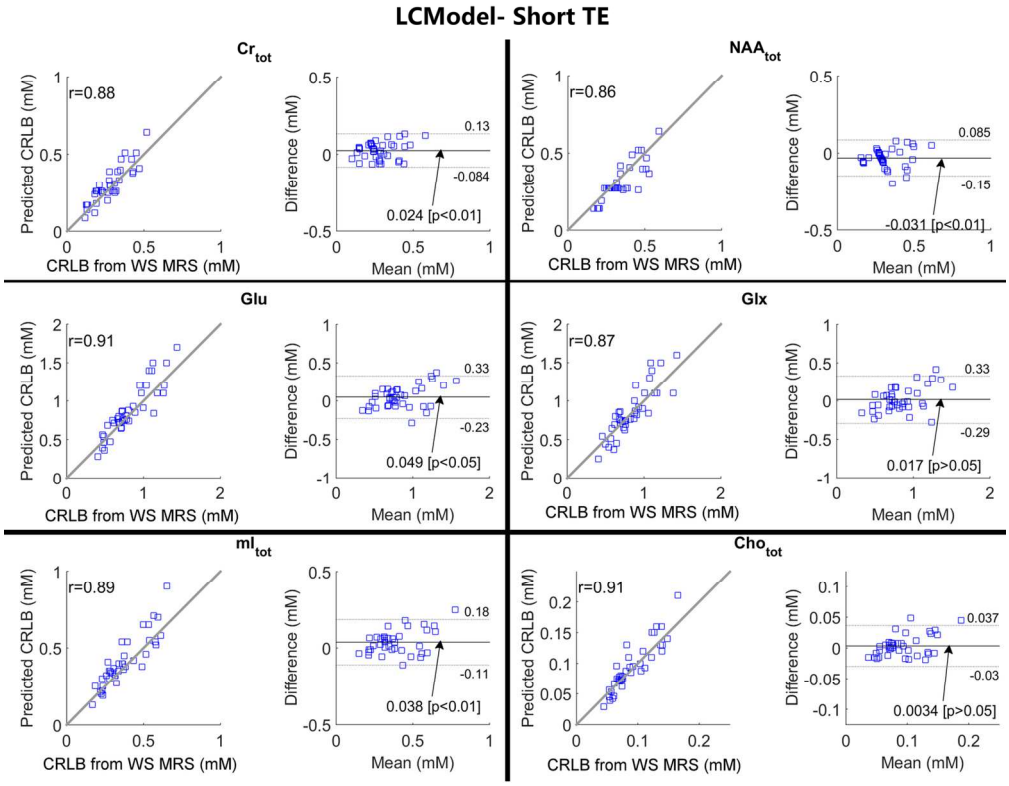

Figure 6: Illustration of the in vivo performance of the proposed error forecast method for the case of LCModel at short TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 6
140x112mm (300 x 300 DPI)

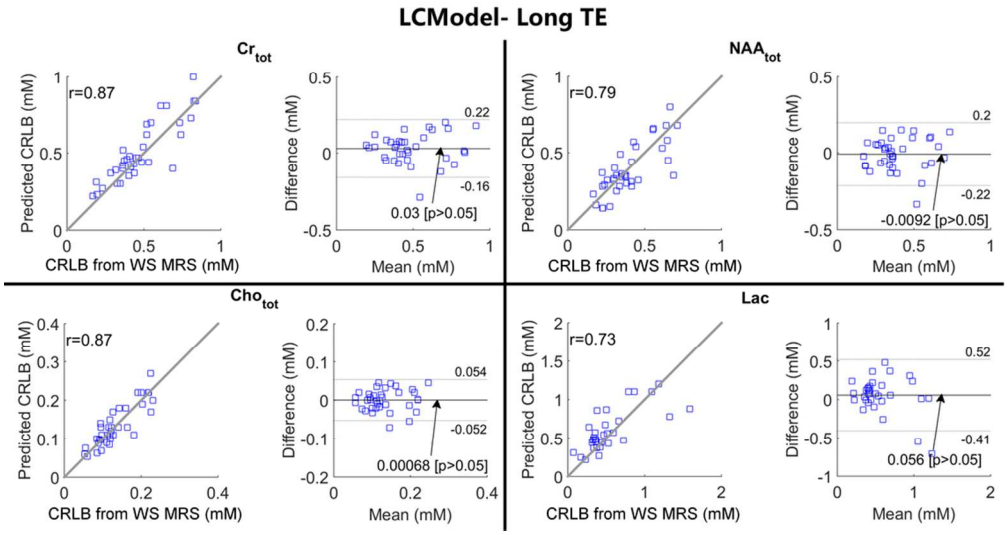

Figure 7: Illustration of the in vivo performance of the proposed error forecast method for the case of LCModel at long TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 7
96x52mm (300 x 300 DPI)

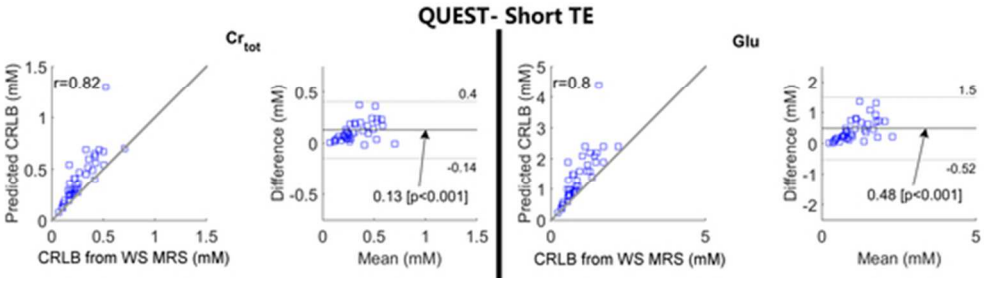

Figure 8: Illustration of the in vivo performance of the proposed error forecast method for the case of QUEST at short TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 8
49x14mm (300 x 300 DPI)

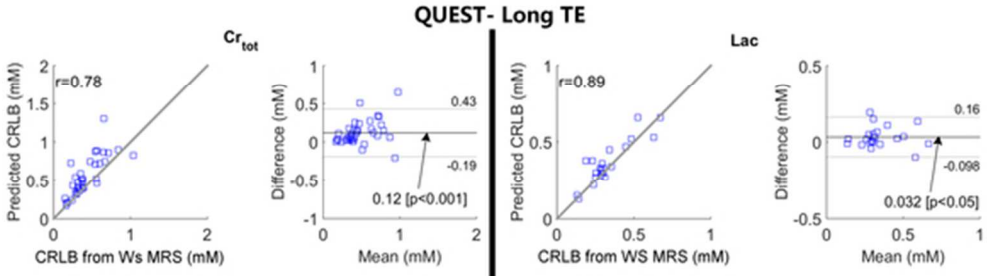

Figure 9: Illustration of the in vivo performance of the proposed error forecast method for the case of QUEST at long TE. The graphs on the left show the correlation between the forecast error using the single water shot and the CRLB as determined from fitting the corresponding water-suppressed spectra, while on the right the same data is presented in the form of Bland-Altman plots. The gray line on the correlation plot is the identity line. The Bland-Altman plots include 95 % confidence intervals for the differences and the significance values for these offsets.

Figure 9
49x14mm (300 x 300 DPI)